

Ecological proteomics: is the field ripe for integrating proteomics into evolutionary ecology research?



"*Nothing in biology makes sense except in the light of evolution*", the title of a talk that Theodosius G. Dobzhansky (January 25, 1900 – December 18, 1975) prominent geneticist and evolutionary biologist delivered to the American Association of Biology Teachers [3] remains highly relevant more than 40 years after its formulation. Dobzhansky was one of the principal architects of the so-called synthetic theory of evolution [1], a theoretical framework integrating Mendelian genetics with Darwin's theory of evolution [2]. Dobzhansky developed the synthetic theory beyond the early mathematical models of RA Fisher, S Wright, JBS Haldane, and SS Chetverikov showing that Mendelian genetics was completely consistent with the idea of evolution driven by natural selection, and produced a comprehensive evolutionary theory that carried population genetics forward to explain the process of speciation [4]. Darwin in his theory of evolution [5] was not aware of Mendel's work and the later synthetic theory regards heritable variation as arising randomly, irrespective of the current adaptive needs imposed by the environment. It is also well established that natural selection produces highly structured patterns of phenotypic variation, which differ among species. The paradigm of 'move, adapt or die', defines the ways in which organisms respond to environmental stressors [6]. Adaptive radiation involves the diversification of a lineage into species that differ in phenotypic traits that can be used to exploit different environmental resources. Ecological opportunities also arise when new or previously inaccessible resources are encountered or are newly exploitable following the acquisition of an evolutionary innovation [7]. Alternatively, populations of the same species might also evolve independently and differentiate by genetic drift, leading to reproductive isolation and ultimately to the formation of new species [7–9]. Regardless of the actual evolutionary scenario, for natural selection to shape the properties of populations in their natural environment, genetic variation (originated randomly by mutation), or key genes with high plastic capacity (leading to plastic phenotypes), need to be present in the population [8,10].

Understanding the molecular basis underlying the effects of genetic variation on phenotype, performance and fitness demands qualitative and quantitative comparisons of the temporal and spatial patterns of genetic and phenotypic variability of adaptive traits [11]. The value of examining a biological system at multiple levels of organization has been summarized by the American biologist and ecological physiologist George A. Bartholomew, who in 1966 wrote "*there are a number of levels of biological integration and... each level finds its explanations of mechanism in the levels below and significance in the levels above it*" [12]. Identifying the molecular basis of adaptations in natural populations has been recognized as an important yet largely unrealized goal in evolutionary biology [13]. An important task in evolutionary ecology is to

identify genes that matter for the processes of adaptation, and ultimately speciation, and to understand the relative importance of structural changes in proteins versus gene regulation effects as determinants of adaptive variation in phenotype [14]. Such information is of broad significance because it bears on fundamental questions about the connection between genotype and phenotype for fitness-related traits. There are a number of examples of this link in a variety of organisms in which structural changes have been quantified by identifying mutations in protein coding regions of genes that have functionally verified effects on adaptive variation at the phenotype level [14]. Likewise, gene regulation effects on adaptive variation are increasingly documented [15,16]. The importance of gene regulation in natural populations has been inferred from measurements of variation in gene expression using a variety of techniques [17], the most widely approach being the measurement of variation in mRNA levels using next-generation sequencing techniques or microarrays developed for model organisms [15]. While this approach is powerful, allowing the simultaneous assessment of genome-scale gene transcription variation, it makes the assumption that there is a close link between transcript abundance and protein levels. However, applying these approaches to non-model organisms can represent a formidable challenge. Genotyping-by-sequencing (GBS) and restriction-site-associated DNA sequencing (RAD) approaches may alleviate this situation by enabling researchers to identify specific genomic regions that may have experienced natural selection [18] (but read also [19] for a critical assessment of the utility, technical complexity and sources of bias for different RAD protocols).

Developments in transcriptome and proteome analysis have progressed on parallel tracks. Both technological platforms stand out in their nominal capacity to measure, with high coverage and untargeted fashion, the entire set of mRNAs/proteins expressed by a complex biological system (e.g. [20–23]). Although global analysis of mRNA abundance has become the standard approach for exploring differences in the genes that matter to organisms undergoing environmental stress, adaptive phenotypic trait variation is often controlled by tissue-specific postgenomic mechanisms affecting protein abundance [24–26]. Hence, the proteome represents a closer proxy to the phenotype, on which selection acts, than the transcriptome [27]. In this regard, it is worth mentioning the premonitory work published long before the advent of the -omics era by King and Wilson [28], which upon comparing the human and chimpanzee genes and proteins concluded that "*all the biochemical methods agree that the genetic distance between humans and the chimpanzees is probably too small to account for their substantial organismal differences*", and that "*evolutionary changes in anatomy and way of life are more often based on changes in the mechanisms controlling*

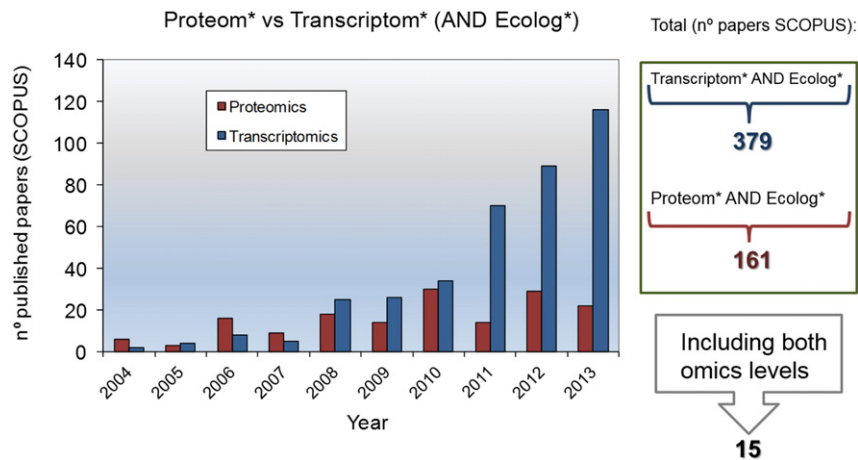


Fig. 1. Proteomic vs. transcriptomic papers on evolutionary/molecular ecology studies (source: SCOPUS). The bar histogram displays the percentages of proteomic (red bar) and transcriptomic (blue bar) studies matching the keywords “proteome*” AND “ecolog*” and “transcriptom*” AND “ecolog*”, respectively. The raw results were manually revised by checking that the focus of each publication was, either directly or indirectly, a truly evolutionary/molecular ecology question. Total number of publications that met this last criterion for each search category (or both) are displayed on the left.

the expression of genes than on the sequence changes in proteins”. Although this example is an extreme case, it serves to remind us that unlike the genome, transcriptomes and proteomes are dynamic and variation can be found in different cells, tissues, organs and developmental stages. In addition, there is increasing evidence supporting the view that only a low number of genetic changes, possibly involving regulatory genes, produce a larger number of adaptive phenotypic changes in a coordinated way [29]. It follows that understanding adaptation in changing environments requires consideration of the pleiotropic effects of specific adaptive mutations, which is the phenomenon that a mutation at one gene can affect more than one phenotypic character. In other words, understanding how genetic variation is translated into phenotypic variation involves detecting both loci affecting the phenotypic association between characters and the patterns of pleiotropic variation in the phenotype.

The proteome can be regarded as a molecular phenotype, and the expression of individual proteins can be treated as quantitative genetic characters [30]. However, despite the importance of examining protein expression for assessing how the underlying phenotypic variation is influenced by selection [31], the investigation of quantitative trait loci (QTL), mRNA expression (eQTL) and protein abundance (pQTL), using combined transcriptomic and proteomic analyses, has been only occasionally included in evolutionary ecology studies [32]. Fig. 1 displays the result of a literature survey in SCOPUS for molecular ecology papers where a transcriptomic and/or a proteomic approach was used. The combined use of both -omic technologies was employed in only 15 studies, mainly dealing with model microorganisms and eukaryotic organisms with fully sequenced genomes. Likewise, only a few presentations at the recent annual meeting of the Society for Integrative and Comparative Biology (Portland, OR, 3-7 January 2016) included proteomic results. A likely reason for the low use of proteomics in evolutionary ecology studies lies in the fact that, until recently, techniques for large-scale measurements of protein abundance lag far behind those for measuring transcript abundance. This circumstance made the application of quantitative proteomic approaches in non-model organisms, albeit representing the vast majority of the biodiversity on Earth [33], a formidable challenge. This situation has changed in recent years [34], and the application of high-throughput proteomics to detect proteins involved in local adaptation in species with limited genomic resources is gaining momentum [11,35,36] (see also reviews [37,38] in this issue).

Ecological proteomics of non-model organisms, which are often characterized by unique genomic and phenotypic characteristics, offers an exciting and complementary reverse genetics approach towards the study of adaptation, trait evolution, and species divergence, including

the identification of ontogenetic changes and the impact of non-genetic mechanisms in the evolution of phenotypic diversity [39]. Another argument for incorporating proteomics analyses in evolutionary ecology studies is the possibility of confirming and revising existing genome annotations [40] and discovering completely new genes [41]. An example is the discovery using a shotgun proteomic of 19 previously unannotated genes encoding seminal fluid proteins (Sfps) that are transferred from males to females during mating in *Drosophila* [42]. Despite the paramount importance of sexual reproduction and reproductive isolation mechanisms in speciation, molecular studies on gametes are scarce. The molecular basis of fertilisation, including the basis of sperm-egg recognition system, is still a poorly understood basic process even for mammalian-model organisms [43]. This Special Issue of the Journal of Proteomics brings together a number of proteomics studies carried out mainly in non-model organisms, and includes several papers addressing proteomic studies of reproductive systems. These, and the disparate collection of papers collected in this Special Issue, highlight the high potential of ecological proteomics to address fundamental questions in ecology and evolutionary biology, allowing researchers to directly link environmental conditions to changes in protein expression. Is the field ripe for integrating proteomics into evolutionary ecology research?

References

- [1] F.J. Ayala, Theodosius Dobzhansky: the man and the scientist, *Annu. Rev. Genet.* 10 (1976) 1–6.
- [2] T. Dobzhansky, *Genetics and the Origin of Species*, Columbia University Press, New York, 1937.
- [3] T. Dobzhansky, Nothing in biology makes sense except in the light of evolution, *Am. Biol. Teach.* 35 (1973) 125–129.
- [4] E. Mayr, *Speciational Evolution or Punctuated Equilibria*, in: S. Peterson (Ed.), *The Dynamics of Evolution* (Somit A, Cornell University Press, New York 1992, pp. 21–48.
- [5] C. Darwin, *On the origin of species*, John Murray, London, UK, 1859.
- [6] D. Porcelli, R.K. Butlin, K.J. Gaston, D. Joly, R.R. Snook, The environmental genomics of metazoan thermal adaptation, *Heredity* 114 (2015) 502–514.
- [7] J.B. Yoder, E. Clancey, S. Des Roches, J.M. Eastman, L. Gentry, W. Godsoe, T.J. Hagey, D. Jochimsen, B.P. Oswald, J. Robertson, B.A. Sarver, J.J. Schenk, S.F. Spear, L.J. Harmon, Ecological opportunity and the origin of adaptive radiations, *J. Evol. Biol.* 23 (2010) 1581–1596.
- [8] D. Schluter, Evidence for ecological speciation and its alternative, *Science* 323 (2009) 737–741.
- [9] M. Kimura, *The neutral theory of molecular evolution*, Cambridge University Press, Cambridge, 1983.
- [10] D.I. Dayan, D.L. Crawford, M.F. Oleksiak, Phenotypic plasticity in gene expression contributes to divergence of locally adapted populations of *Fundulus heteroclitus*, *Mol. Ecol.* 24 (2015) 3345–3359.

- [11] A.C. Dalziel, S.M. Rogers, P.M. Schulte, Linking genotypes to phenotypes and fitness: how mechanistic biology can inform molecular ecology, *Mol. Ecol.* 18 (2009) 4997–5017.
- [12] G.A. Bartholomew, Interaction of physiology and behavior under natural conditions, in: Bowman RI (Ed.), *The Galapagos: Proceedings of the Symposia of the Galapagos International Scientific Project*, University of California Press, Berkeley, University Extension, National Science Foundation (U.S.) 1966, pp. 39–45.
- [13] G.B. Golding, A.M. Dean, The structural basis of molecular adaptation, *Mol. Biol. Evol.* 15 (1998) 355–369.
- [14] H. Hoekstra, J. Coyne, The locus of evolution: evo-devo and the genetics of adaptation, *Evolution* 61 (2007) 995–1016.
- [15] A. Whitehead, D.L. Crawford, Variation within and among species in gene expression: raw material for evolution, *Mol. Ecol.* 15 (2006) 1197–1211.
- [16] H.L. Gibbs, L. Sanz, J.J. Calvete, Snake population venomomics: proteomics-based analyses of individual variation reveals significant gene regulation effects on venom protein expression in *Sistrurus rattlesnakes*, *J. Mol. Evol.* 68 (2009) 113–125.
- [17] G.A. Wray, M.W. Hahn, E. Abouheif, J.P. Balhoff, M. Pizer, M.V. Rockman, L.A. Romano, The evolution of transcriptional regulation in eukaryotes, *Mol. Biol. Evol.* 20 (2003) 1377–1419.
- [18] S.R. Narum, C.A. Buerkle, J.W. Davey, M.R. Miller, P.A. Hohenlohe, Genotyping-by-sequencing in ecological and conservation genomics, *Mol. Ecol.* 22 (2013) 2841–2847.
- [19] J.B. Puritz, M.V. Matz, R.J. Toonen, J.N. Weber, D.I. Bolnick, C.E. Bird, Demystifying the RAD fad, *Mol. Ecol.* 23 (2014) 5937–5942.
- [20] D.R. Rokyt, A.R. Lemmon, M.J. Margres, K. Aronow, The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*), *BMC Genomics* 13 (2012) 312.
- [21] S. Eichberg, L. Sanz, J.J. Calvete, D. Pla, Constructing comprehensive venom proteome reference maps for integrative venomomics, *Expert Rev. Proteomics* 12 (2015) 557–573.
- [22] S. Papakostas, A. Vasemägi, J.P. Vähä, M. Himberg, L. Peil, C.R. Primmer, A proteomics approach reveals divergent molecular responses to salinity in populations of European whitefish (*Coregonus lavaretus*), *Mol. Ecol.* 21 (2012) 3516–3530.
- [23] E.J. Foss, D. Radulovic, S.A. Shaffer, D.M. Ruderfer, A. Bedalov, D.R. Goodlett, L. Kruglyak, Genetic basis of proteome variation in yeast, *Nat. Genet.* 39 (2007) 1369–1375.
- [24] A.L. Bauernfeind, E.J. Soderblom, M.E. Turner, M.A. Moseley, J.J. Ely, P.R. Hof, C.C. Sherwood, G.A. Wray, C.C. Babbitt, Evolutionary Divergence of Gene and Protein Expression in the Brains of Humans and Chimpanzees, *Genome Biol. Evol.* 7 (2015) 2276–2288.
- [25] J. Durban, A. Pérez, L. Sanz, A. Gómez, F. Bonilla, S. Rodríguez, D. Chacón, M. Sasa, Y. Angulo, J.M. Gutiérrez, J.J. Calvete, Integrated "omics" profiling indicates that miRNAs are modulators of the ontogenetic venom composition shift in the Central American rattlesnake, *Crotalus simus simus*, *BMC Genomics* 14 (2013) 234.
- [26] N.R. Casewell, S.C. Wagstaff, W. Wüster, D.A. Cook, F.M. Bolton, S.I. King, D. Pla, L. Sanz, J.J. Calvete, R.A. Harrison, Medically important differences in snake venom composition are dictated by distinct postgenomic mechanisms, *Proc. Natl. Acad. Sci. U. S. A.* 111 (2014) 9205–9210.
- [27] M.E. Feder, J.C. Walser, The biological limitations of transcriptomics in elucidating stress and stress responses, *J. Evol. Biol.* 18 (2005) 901–910.
- [28] M.C. King, A.C. Wilson, Evolution at two levels in humans and chimpanzees, *Science* 188 (1975) 107–116.
- [29] C.G. Knight, N. Zitzmann, S. Prabhakar, R. Antrobus, R. Dwek, H. Hebestreit, P.B. Rainey, Unraveling adaptive evolution: how a single point mutation affects the protein coregulation network, *Nat. Genet.* 38 (2006) 1015–1022.
- [30] A.P. Diz, D.O.F. Skibinski, Evolution of 2-DE protein patterns in a mussel hybrid zone, *Proteomics* 7 (2007) 2111–2120.
- [31] A. Vasemägi, C.R. Primmer, Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies, *Mol. Ecol.* 14 (2005) 3623–3642.
- [32] A.P. Diz, M. Martínez-Fernández, E. Rolán-Alvarez, (2012) Proteomics in evolutionary ecology: linking the genotype with the phenotype, *Mol. Ecol.* 21 (2012) 1060–1080.
- [33] S.B. Hedges, The origin and evolution of model organisms, *Nat. Rev. Genet.* 3 (2002) 838–849.
- [34] J.J. Calvete, Challenges and prospects of proteomics of non-model organisms, *J. Proteome* 105 (2014) 1–4.
- [35] S. Papakostas, A. Vasemägi, M. Himberg, C.R. Primmer, Proteome variance differences within populations of European whitefish (*Coregonus lavaretus*) originating from contrasting salinity environments, *J. Proteome* 105 (2014) 144–150.
- [36] A. Hidalgo-Galiana, M. Monge, D.G. Biron, F. Canals, I. Ribera, A. Cieslak, Reproducibility and consistency of proteomic experiments on natural populations of a non-model aquatic insect, *PLoS One* 9 (2014), e104734.
- [37] B. Baer, A.H. Millar, Proteomics in evolutionary ecology, *J. Proteome* 135 (2016) 4–11 (this issue).
- [38] K. Sunagar, D. Morgenstern, A.M. Reitzel, Y. Moran, Ecological venomomics: How genomics, transcriptomics and proteomics can shed new light on the ecology and evolution of venom, *J. Proteome* 135 (2016) 62–72 (this issue).
- [39] J.J. Calvete, Snake venomomics: from the inventory of toxins to biology, *Toxicon* 75 (2013) 44–62.
- [40] J. Armengaud, J. Trapp, O. Pible, O. Geffard, A. Chaumot, E.M. Hartmann, Non-model organisms, a species endangered by proteogenomics, *J. Proteome* 105 (2014) 5–18.
- [41] G.D. Findlay, W.J. Swanson, Proteomics enhances evolutionary and functional analysis of reproductive proteins, *BioEssays* 32 (2010) 26–36.
- [42] G.D. Findlay, M.J. MacCoss, W.J. Swanson, Proteomic discovery of previously unannotated, rapidly evolving seminal fluid genes in *Drosophila*, *Genome Res.* 19 (2009) 886–896.
- [43] E. Bianchi, B. Doe, D. Goulding, G. Wright, Juno is the egg Izumo receptor and is essential for mammalian fertilization, *Nature* 508 (2014) 483–487.

Angel P. Diz

Department of Biochemistry, Genetics and Immunology, Faculty of Biology,
University of Vigo, Vigo, Spain
E-mail address: angel.p.diz@uvigo.es.

Juan J. Calvete

Instituto de Biomedicina de Valencia, CSIC, Valencia, (Spain)
E-mail address: jcalvete@ibv.csic.es.